

04C0

02-13-01

OIPE

#3

RAW SEQUENCE LISTING                      DATE: 01/29/2001  
PATENT APPLICATION: US/09/759,207              TIME: 13:14:20

Input Set : A:\Seq.txt  
Output Set: N:\CRF3\01292001\I759207.raw

## SEQUENCE LISTING

3 (1) GENERAL INFORMATION:  
C--> 4 (i) APPLICANT: Iris Pecker et al.  
5 (ii) TITLE OF INVENTION: HEPARANASE SPECIFIC MOLECULAR PROBES  
6 AND THEIR USE IN RESEARCH AND MEDICAL  
7 APPLICATIONS  
8 (iii) NUMBER OF SEQUENCES: 7  
9 (iv) CORRESPONDENCE ADDRESS:  
10 (A) ADDRESSEE: G. E. Ehrlich (1995) Ltd.  
11 c/o Anthony Castorina  
12 (B) STREET: 2001 Jefferson Davis Highway, Suite 207  
13 (C) CITY: Arlington  
14 (D) STATE: Virginia  
15 (E) COUNTRY: United States of America  
16 (F) ZIP: 22202  
17 (v) COMPUTER READABLE FORM:  
18 (A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
19 (B) COMPUTER: Twinhead\* Slimnote-890TX  
20 (C) OPERATING SYSTEM: MS DOS version 6.2,  
21 Windows version 3.11  
22 (D) SOFTWARE: Word for Windows version 2.0 converted to  
23 an ASCII file  
24 (vi) CURRENT APPLICATION DATA:  
C--> 25 (A) APPLICATION NUMBER: US/09/759,207  
C--> 26 (B) FILING DATE: 16-Jan-2001  
27 (C) CLASSIFICATION:  
28 (vii) PRIOR APPLICATION DATA:  
29 (A) APPLICATION NUMBER: 08/922,180  
30 (B) FILING DATE: September 2, 1997  
31 (A) APPLICATION NUMBER: 09/071,739  
32 (B) FILING DATE: May 1, 1998  
33 (A) APPLICATION NUMBER: 09/322,977  
34 (B) FILING DATE: June 1, 1999  
35 (viii) ATTORNEY/AGENT INFORMATION:  
36 (A) NAME: Sol Sheinbein  
37 (B) REGISTRATION NUMBER: 25,457  
38 (C) REFERENCE/DOCKET NUMBER: 00/21505  
39 (ix) TELECOMMUNICATION INFORMATION:  
40 (A) TELEPHONE: 972-3-6127676  
41 (B) TELEFAX: 972-3-6127575  
42 (C) TELEX:  
44 (2) INFORMATION FOR SEQ ID NO: 1:  
45 (i) SEQUENCE CHARACTERISTICS:  
46 (A) LENGTH: 1721  
47 (B) TYPE: nucleic acid  
48 (C) STRANDEDNESS: double  
49 (D) TOPOLOGY: linear

ENTERED

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/759,207  
 DATE: 01/29/2001  
 TIME: 13:14:20

Input Set : A:\Seq.txt  
 Output Set: N:\CRF3\01292001\I759207.raw

C--> 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 51 CTAGAGCTTT CGACTCTCCG CTGCGCGGCA GCTGGCGGGG GGAGCAGCCA GGTGAGCCCA 60  
 52 AGATGCTGCT GCGCTCGAAG CCTGCGCTGC CGCCGCCGCT GATGCTGCTG CTCTGGGGG 120  
 53 CGCTGGGTCC CCTCTCCCTT GCGCGCTGC CCCGACCTGC GCAAGCACAG GACGTCGTGG 180  
 C--> 54 ACCTGGACTT CTTACCCAG GAGCCGCTGC ACCTGGTGAG CCCCTCGTTC CTGTCCGTCA 240  
 55 CCATTGACGC CAACCTGGCC ACCGACCCGC GGTTCCTCAT CCCTCTGGGT TCTCCAAAGC 300  
 56 TTCGTACCTT GGCAGAGGC TTGTCTCCTG GTTACCTGAG GTTGGTGAG ACCAAGACAG 360  
 57 ACTTCCTAAT TTPCGATCCC AAGAAGGAAT CAACCTTTGA AGAGAGAAGT TACTGGCAAT 420  
 58 CTCAGTCAA CCAGGATATT TGCAAAATATG GATCCATCCC TCCTGATGTG GAGGAGAAGT 480  
 59 TACGGTTGGA ATGGCCCTAC CAGGAGCAAT TGCTACTCCG AGAACACTAC CAGAAAAGT 540  
 60 TCAAGAACAG CACCTACTCA AGAAGCTCTG TAGATGTGCT ATACACTTTT GCAAACTGCT 600  
 61 CAGGACTGGA CTTGACTTTT GGCCTAAATG CGTTATTAAG AACAGCAGAT TTGCAGTGA 660  
 62 ACAGTTCTAA TGCTCAGTTG CTCTGGACT ACTGCTCTTC CAAGGGGTAT AACATTTCTT 720  
 63 GGGAACTAGG CAATGACCTT AACAGTTTCC TTAAGAAGGC TGATATTTTC ATCAATGGGT 780  
 64 CGCAGTTAGG AGAAGATTAT ATTCAATTGC ATAACTTCTT AAGAAAGTCC ACCTTCAAAA 840  
 65 ATGCAAAACT CTATGGTCTT GATGTTGGTC AGCCTCGAAG AAAGACGGCT AAGATGCTGA 900  
 66 AGAGCTTCCT GAAGGCTGGT GGAGAAGTGA TTGATTGAGT TACATGGCAT CACTACTATT 960  
 67 TGAATGGAGC GACTGCTACC AGGGAAGATT TTCTAAACCC TGATGATTG GACATTTTTA 1020  
 68 TTTCATCTGT GCAAAAAGTT TTCCAGGTGG TTGAGAGCAC CAGGCTGGC AAGAAGTCT 1080  
 69 GGTTAGGAGA AACAAGCTCT GCATATGGAG GCGGAGCGCC CTTCCTATCC GACACCTTTG 1140  
 70 CAGCTGGCTT TATGTGGCTG GATAAATFGG GCCTGTTCAGC CCGAATGGGA ATAGAAGTGG 1200  
 71 TGATGAGGCA AGTATCTTTT GGAGCAGGAA ACTACCATTT AGTGGATGAA AACTTCGATC 1260  
 72 CTTTACCTGA TTATTGGCTA TCTCTTCTGT TCAAGAAATT GGTGGGCACC AAGGTGTTAA 1320  
 73 TGGCAAGCCT GCAAGTTTCA AAGAGAAGGA AGCTTCGAGT ATACCTTCAT TGCACAAACA 1380  
 74 CTGACAAATC AAGGTATAAA GAAGGAGATT TAACTCTGTA TGCCATAAAC CTCATAACG 1440  
 75 TCACCAAGTA CTTGCGGTTA CCCTATCCTT TTTCTAACAA GCAAGTGGAT AAATACCTTC 1500  
 76 TAAGACCTTT GGGACCTCAT GGATTACTTT CCAATCTGT CCAACTCAAT GGTCTAACTC 1560  
 77 TAAAGATGGT GGATGATCAA ACCTTGCCAC CTTTAATGGA AAAACCTCTC CGGCCAGGAA 1620  
 78 GTTCACTGGG CTTGCCAGCT TTCTCATATA GTTTTTTGT GATAAGAAAT GCCAAAGTTG 1680  
 79 CTGCTTGCTT CTGAAAATAA AATATACTAG TCCTGACACT G 1721  
 81 (2) INFORMATION FOR SEQ ID NO: 2:  
 82 (i) SEQUENCE CHARACTERISTICS:  
 83 (A) LENGTH: 543  
 84 (B) TYPE: amino acid  
 85 (C) STRANDEDNESS: single  
 86 (D) TOPOLOGY: linear  
 C--> 87 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 89 Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu  
 90 5 10 15  
 92 Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro  
 93 20 25 30  
 95 Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln.Glu Pro  
 96 35 40 45  
 98 Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn  
 99 50 55 60  
 101 Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu  
 102 65 70 75 80  
 104 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly  
 105 85 90 95

## RAW SEQUENCE LISTING

DATE: 01/29/2001

PATENT APPLICATION: US/09/759,207

TIME: 13:14:20

Input Set : A:\Seq.txt

Output Set: N:\CRF3\01292001\I759207.raw

```

107 Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe
108      100      105      110
110 Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys
111      115      120      125
113 Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp
114      130      135      140
116 Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe
117 145      150      155      160
119 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
120      165      170      175
122 Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu
123      180      185      190
125 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu
126      195      200      205
128 Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn
129      210      215      220
131 Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser
132 225      230      235      240
134 Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser
135      245      250      255
137 Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg
138      260      265      270
140 Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu
141      275      280      285
143 Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr
144      290      295      300
146 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile
147 305      310      315      320
149 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly
150      325      330      335
152 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala
153      340      345      350
155 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys
156      355      360      365
158 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val
159      370      375      380
161 Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro
162 385      390      395      400
164 Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
165      405      410      415
167 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg
168      420      425      430
170 Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly
171      435      440      445
173 Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu
174      450      455      460
176 Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu
177 465      470      475      480
179 Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn

```

## RAW SEQUENCE LISTING

DATE: 01/29/2001

PATENT APPLICATION: US/09/759,207

TIME: 13:14:20

Input Set : A:\Seq.txt

Output Set: N:\CRF3\01292001\I759207.raw

```

180          485          490          495
182 Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met
183          500          505          510
185 Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser
186          515          520          525
188 Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile
189          530          535          540          543
191 (2) INFORMATION FOR SEQ ID NO: 3:
192 (i) SEQUENCE CHARACTERISTICS:
193 (A) LENGTH: 1721
194 (B) TYPE: Nucleic acid
195 (C) STRANDEDNESS: Double
196 (d) TOPOLOGY: linear
C--> 197 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
199          CT AGA GCT TTC GAC 14
201 TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA GGT GAG CCC AAG 62
203 ATG CTG CTG CGC TCG AAG CCT GCG CTG CCG CCG CTG ATG CTG CTG 110
204 Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu
205          5          10          15
207 CTC CTG GGG CCG CTG GGT CCC CTC TCC CCT GGC GCC CTG CCC CGA CCT 158
208 Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro
209          20          25          30
211 GCG CAA GCA CAG GAC GTC GTG GAC CTG GAC TTC TTC ACC CAG GAG CCG 206
212 Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro
213          35          40          45
215 CTG CAC CTG GTG AGC CCC TCG TTC CTG TCC GTC ACC ATT GAC GCC AAC 254
216 Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn
217          50          55          60
219 CTG GCC ACG GAC CCG CGG TTC CTC ATC CTC CTG GGT TCT CCA AAG CTT 302
220 Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu
221          65          70          75          80
223 CGT ACC TTG GCC AGA GGC TTG TCT CCT GCG TAC CTG AGG TTT GGT GGC 350
224 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly
225          85          90          95
227 ACC AAG ACA GAC TTC CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT 398
228 Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe
229          100          105          110
231 GAA GAG AGA AGT TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA 446
232 Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys
233          115          120          125
235 TAT GGA TCC ATC CCT CCT GAT GTG GAG GAG AAG TTA CCG TTG GAA TGG 494
236 Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp
237          130          135          140
239 CCC TAC CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC 542
240 Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe
241          145          150          155          160
243 AAG AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 590
244 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
245          165          170          175

```

## RAW SEQUENCE LISTING

DATE: 01/29/2001

PATENT APPLICATION: US/09/759,207

TIME: 13:14:20

Input Set : A:\Seq.txt

Output Set: N:\CRF3\01292001\I759207.raw

```

247 GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCC TTA TTA 638
248 Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu
249 180 185 190
251 AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG CTC CTG 686
252 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu
253 195 200 205
255 GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA CTA GGC AAT 734
256 Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn
257 210 215 220
259 GAA CCT AAC AGT TTC CTT AAG AAG GCT GAT ATT TTC ATC AAT GGG TCG 782
260 Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser
261 225 230 235 240
263 CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA CTT CTA AGA AAG TCC 830
264 Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser
265 245 250 255
267 ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT GAT GTT GGT CAG CCT CGA 878
268 Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg
269 260 265 270
271 AGA AAG ACG GCT AAG ATG CTG AAG AGC TTC CTG AAG GCT GGT GGA GAA 926
272 Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu
273 275 280 285
275 GTG ATT GAT TCA GTT ACA TGG CAT CAC TAC TAT TTG AAT GGA CGG ACT 974
276 Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr
277 290 295 300
279 GCT ACC AGG GAA GAT TTT CTA AAC CCT GAT GTA TTG GAC ATT TTT ATT 1022
280 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile
281 305 310 315 320
283 TCA TCT GTG CAA AAA GTT TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC 1070
284 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly
285 325 330 335
287 AAG AAG GTC TCG TTA GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG 1118
288 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala
289 340 345 350
291 CCC TTG CTA TCC GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA 1166
292 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys
293 355 360 365
295 TTG GGC CTG TCA GCC CGA ATG GGA ATA GAA GTG GTG ATG AGG CAA GTA 1214
296 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val
297 370 375 380
299 TTC TTT GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT 1262
300 Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro
301 385 390 395 400
303 TTA CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1310
304 Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
305 405 410 415
307 AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT CGA 1358
308 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg
309 420 425 430
311 GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA GAA GGA 1406

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/759,207

DATE: 01/29/2001

TIME: 13:14:21

Input Set : A:\Seq.txt

Output Set: N:\CRF3\01292001\I759207.raw

L:4 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]  
L:25 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:26 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:50 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:54 M:111 C: (47) String data converted to upper case,  
L:87 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:197 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:347 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:356 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:365 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:374 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]